Journal of Choice Modelling, 5(2), pp 81-97
www.jocm.org.uk

# Design effects in a meta-analysis of river health choice experiments in Australia

John Rolfe[1,*] Roy Brouwer[2,†]

[1]CQUniversity, Centre for Environmental Management, Bruce Highway, Rockhampton, Queensland 4702, Australia.
[2]Department of Environmental Economics, Institute for Environmental Studies, VU University Amsterdam, De Boelelaan 1087, 1081 HV Amsterdam, The Netherlands.

Received 4 July 2011, revised version received 14 October 2011, accepted 22 June 2012

## Abstract

While meta-analysis is typically used to identify value estimates for benefit transfer, applications also provide insights into the potential influence of methodological design characteristics on results of non-market valuation experiments. In this paper, a meta-analysis of nineteen choice modeling (CM) studies in Australia is conducted generating 145 individual value estimates relating to river health. Implicit prices of different measures and scales of river health were transformed into a common standard of willingness to pay (WTP) per kilometer of river in good health. A Tobit model was used to identify the relationships between this dependent variable and a large number of study design characteristics. While there is evidence that the dimensions of choice tasks and description of attributes influence value estimates, there is also evidence that the way tradeoffs and payment mechanisms are framed are equally important. The results of this meta-analysis suggest that more attention should be paid to the way tradeoffs are framed in choice experiments relative to internal choice set structure and data analysis.

*Keywords: Meta analysis, Choice modelling, River health*

* Corresponding author, T: +61-7-49232132, F: + 61-7-49309159, j.rolfe@cqu.edu.au
† T: +31-20-5989 5608, F: + 31-20-5989 553, r.brouwer@vu.nl



ISSN 1755-5345

# 1 Introduction

Benefit transfer is an attractive alternative to original valuation research as results from a small number of targeted studies are transferred to other policy contexts and sites of interest, providing a cost-effective means of extending economic analysis (Brookshire and Neil 1992, Brouwer 2000, Rolfe and Bennett 2006, Rosenberger and Stanley 2006, Johnston and Rosenberger 2010). However, there are a number of questions about the validity and accuracy of benefit transfer applications, with concerns that large transfer errors may limit the usefulness of results (Brouwer and Spaninks 1999, Brouwer and Bateman 2005, Rosenberger and Stanley 2006, Rolfe 2006, Johnston and Rosenberger 2010, Bateman et al. 2011). Meta-analysis, defined as the systematic quantitative summary of evidence across empirical studies (Glass et al. 1981), has been advanced as a way of generating more robust and reliable estimates of values for use in benefit transfer (Bateman and Jones 2003, Bergstrom and Taylor 2006, Nelson and Kennedy 2009).

Key advantages of adopting a meta-analysis approach to benefit transfer are that (a) the results of multiple studies can be incorporated, (b) it is possible to control for effects such as sample size and heterogeneity, (c) methodological differences can be identified, and (d) subsequent value functions can be used to predict values for potential target sites (Rosenberger and Loomis 2000, Nelson and Kennedy 2009). Meta-analysis has been widely applied across a range of discipline areas and policy issues (Nelson and Kennedy 2009), including environmental applications such as wetlands (Brouwer et al. 1999, Woodward and Wui 2001, Brander et al. 2006), water (Johnston et al. 2003, van Houtven et al. 2007), and forests (Lindhjem 2007, Zandersen and Tol 2009, Barrio and Loureiro 2010).

Most meta-analysis studies in the field of environmental economics focus on a particular issue or amenity of interest, and draw together results of different studies for analysis, including studies from different methodological frameworks (Van den Bergh et al. 1997). However this type of analysis is limited, for example, where meta-analysis studies incorporate results from different stated preference and revealed preference techniques (Bateman and Jones 2003, Nelson and Kennedy 2009, Johnston and Rosenberger 2010), leading to criticisms that values may not be commensurable (Smith and Pattanayak 2002, Bergstrom and Taylor 2006). A key limitation of such diverse pooling of values is that methodological influences on values, such as the design characteristics of stated preference experiments, may not be distinguishable.

The research reported in this paper involved a meta-analysis where only choice experiments have been included, the first that the authors have been able to identify[1]. The main purpose of the study was to identify if a number of choice experiment related methodological and design characteristics have a significant influence on values, pinpointing areas where analysts should exercise more caution in both experiment designs and the subsequent benefit transfer process. The paper is organized as follows. The following section sets out the relevant meta-analysis methodology. In Section 3, an overview is provided of the available choice experiment literature related to healthy waterways in Australia. A discussion of the key factors that might influence river protection values is provided in section 4, and results of the meta-analysis provided in sections 5 and 6. Final conclusions follow in section 7.

[1] The study presented here is based on the work of Brouwer (2009). Other meta-analysis studies such as Johnston et al. (2005) or Lindhjem (2007) incorporate CE studies along with results from other non-market valuation techniques.

## 2 Meta-analysis methodology

The meta-model used to predict the marginal rate of substitution between income and values for an environmental improvement ($\beta_i$) can be described generally as follows:

$$y_i = \beta_0 + \beta_1 X_{i1} + \ldots + \beta_k X_{ik} + \varepsilon_i \tag{1}$$

where $y_i$ is the estimate of value for the environmental change (usually standardized to a per unit change across studies), $\beta_0$ represents an intercept term, $\beta_i$ are the coefficient terms associated with $X_{ik}$ study variables and moderators, and $\varepsilon_i$ is an error term. Regression analysis is typically employed in meta-analysis, with specialised variations used to address methodological and econometric issues (Nelson and Kennedy 2009).

While interest in meta-analysis has been driven by the potential to improve the accuracy of value transfers, a number of challenges exist in performing a successful analysis (Johnston and Rosenberger 2010). These include the difficulties of ensuring commensurability across data sets, the variation in methods and approaches because the experiments are not controlled, limited data sets and inadequate methods of analysis (Smith and Pattanayak 2002, Johnston and Rosenberger 2010). The choice of a meta-regression model has to take account of data heterogeneity, heteroscedasticity and correlated observations (Nelson and Kennedy 2009). Heterogeneity in primary data occurs because studies have different characteristics and are based on different populations, and because of variations in study designs and methods. Standard approaches to address heterogeneity are to include suitable regressors in the analysis, or to conduct a series of random draws from the data set using a random-effect-size model (Nelson and Kennedy 2009). Heteroskedasticity occurs when the variances across samples are non-constant, for reasons such as differences in sample size and estimation procedures. A standard approach to addressing heteroskedasticity is to weight the observations, preferably with greater focus on observations with lower variance (Nelson and Kennedy 2009). Correlation within and between primary studies can occur when a number of split-samples is generated from a single study, or when the same data set is used for more than one prediction model. Methods to address this include selecting only one sample from each study, using hierarchical regression techniques, or applying mixed fixed and random effects panel data corrections (Nelson and Kennedy 2009).

For this study a mixed effects tobit regression function is applied, given the censored nature of the data (positive WTP values only) and heteroscedasticity (intra-study effects due to similar design). The meta-model used to predict the marginal rate of substitution between income and a healthy waterways attribute (implicit price) can be described more generally as follows:

$$MWTP_i = \beta_j X_{ij} = \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i \tag{2}$$

where $MWTP_i$ is the vector containing the implicit price found in study *i* and $X_{ij}$ represents the design matrix for the covariates, consisting of amenity characteristics (measured through the vector $\beta_1$), population characteristics (measured through the vector $\beta_2$), and methodological study design characteristics (measured through the vector $\beta_3$), with the latter capturing variations in trade-off framing, payment mechanisms, data collection, level description, choice set design, and data analysis. In order to account for heteroscedasticity, i.e., cross-sectional correlations between multiple observations from the same or different studies, model variance is made a function of the covariates. Making the errors depend on explanatory factors and including them in the random part of the model allows a random effects Tobit model to be obtained where the error becomes a composite matrix including the stochastic disturbances associated with the fixed and random effects in the model's design matrix:

$$\mathrm{MWTP}_i^* = \beta_j X_{ij} + \omega_{ij} \tag{3}$$

where $\omega_{ij} = \varepsilon_i + u_{ij}$ with the following distributional characteristics:

$$\omega_{ij} \sim N(0, \Sigma); \Sigma = \begin{bmatrix} \sigma_\varepsilon^2 ; \sigma_\varepsilon \sigma_u \\ \sigma_u \sigma_\varepsilon ; \sigma_u^2 \end{bmatrix} \tag{4}$$

$\varepsilon_i$ is the residual associated with the intercept and $u_{ij}$ with the slope parameter $\beta_j$.

## 3 Case studies and WTP variable

Case studies chosen for this meta-analysis involved non-market values of river protection in Australia. Key steps in the initial stages are to define the environmental issue of interest, identify how studies have been selected, and to describe how data has been coded. To generate a theoretically consistent base for data pooling and analysis, only value estimates from choice experiment studies to protect rivers in Australia were chosen. Marginal tradeoffs in terms of implicit prices (also known as part-worths) were chosen as the dependent variable because these avoided scale parameter issues in comparing results, and allowed only values for selected attributes to be reported. Using compensating surplus estimates as the dependent variable was not identified as practical because of the difficulty in establishing future protection scenarios that were consistent across case studies and the variation in attributes between case studies. Moreover, where only unit changes in single attributes are involved, estimates of compensating surplus collapse to implicit prices. As implicit prices are regularly used to conduct benefit transfer tests (e.g., Morrison et al. 2002), there is no theoretical barrier to their use in pooling study results.

The meta-analysis revealed 154 individual value estimates from separate experiments conducted within nineteen separate choice modelling studies across five states and territories (Table 1). These studies were drawn from a range of published and grey literature sources. Each of the studies was available publicly in some format, and each provided implicit prices and enough study details to populate the meta-analysis. Many studies involved multiple split-samples with different results reported for each experiment.

Definition of the dependent variable proved problematic, as many studies involved slightly different aspects of environmental health. Many researchers defined the environmental good in similar terms such as 'waterways in the catchment remaining in good health', 'waterways in good health' and 'healthy waterways' (Table 1).

## Table 1: Overview of studies included in the meta-analysis

| | Authors | Study year | River catchment | State | Split samples | Implicit price (WTP) measured in terms of |
|---|---|---|---|---|---|---|
| 1 | Van Bueren and Bennett (2004) | 2000 | All waterways (not specified) | National, QLD, WA | 6 | $/hh/year per 10 km restored waterway for fishing or swimming |
| 2 | Morrison and Bennett (2004) | 2000 | Bega, Clarence, Georges, Gwydir, Murrumbidgee | NSW | 31 | $/hh/year + one-time-off per % of river covered with healthy native vegetation / per fish species / for fishable/swimmable water quality whole river / per waterbird & other fauna species |
| 3 | Morrison and Bennett (2006) | 2000 | NSW rivers | NSW | 5 | $/hh/year + one-time-off per % of river covered with healthy native vegetation / per fish species / for fishable/swimmable water quality whole river / per waterbird & other fauna species |
| 4 | Rolfe et al. (2002) | 2000 | Fitzroy, Dawson Comet-Nogoa-Mackenzie | QLD | 7 | $/hh/year per km of waterways in the catchment remaining in good health |
| 5 | Rolfe and Windle (2003) | 2001 | Fitzroy | QLD | 3 | $/hh/year per km of waterways in the catchment remaining in good health |
| 6 | Windle and Rolfe (2004) | 2003 | Fitzroy | QLD | 3 | $/hh/year and one-time-off per km of waterways remaining in good health |
| 7 | Windle and Rolfe (2006) | 2005 | Queensland, Fitzroy, MD, Mackay Whitsun. | QLD | 18 | $/hh/year per % of waterways in good health |
| 8 | Kragt et al. (2007) | 2006 | Goulburn | NSW | 16 | $/hh one-time-off per % native fish species and population level / for % of river length with healthy native vegetation / per native waterbird and animal species |
| 9 | Bennett et al. (2008a) | 2006 | Moorabool, Gellibrand, Goulburn | NSW VIC | 12 | $/hh one-time-off per % native fish species and population level / for % of river length with healthy native vegetation / per native waterbird and animal species |
| 10 | Bennett et al (2008b) | 2006 | Murray River | NSW VIC | 5 | $/hh/year per % of pre-European fish numbers / % of healthy flooded vegetation (river red gums) |
| 11 | Rolfe and Bennett (2009) | 2002 | Fitzroy | QLD | 1 | $/hh/year per km of waterways in the catchment remaining in good health |
| 12 | Kragt and Bennett (2009a) | 2008 | Georges | TAS | 3 | $/hh/year per km of river length with healthy native vegetation |
| 13 | Kragt and Bennett (2009b) | 2008 | Georges | TAS | 2 | $/hh/year per km of river length with healthy native vegetation |
| 14 | Kragt and Bennett (2010) | 2009 | Georges | TAS | 4 | $/hh/year per km of river length with healthy native vegetation |
| 15 | Mazur and Bennett (2009) | 2008 | Lachlan, Namoi, Hawkesb.-Nepean | NSW | 7 | $/hh/year per km of healthy waterways |
| 16 | Mazur (2011) | 2008 | Namoi | NSW | 5 | $/hh/year per km of healthy waterways |
| 17 | Mazur and Bennett (2010) | 2008 | Hawkesbury-Nepean | NSW | 7 | $/hh/year per km of healthy waterways |
| 18 | Hatton MacDonald et al. (2010) | 2010 | Murray | NSW VIC, SA | 18 | $/hh/year per % of healthy vegetation along the River Murray |
| 19 | Morrison et al (2010) | 2010 | Murray | NSW, VIC, SA | 1 | $/hh/year per % of healthy vegetation along the River Murray |

ACT = Australian Capital Territory, QLD = Queensland, WA = Western Australia, NSW = New South Wales, VIC = Victoria, TAS = Tasmania, MD=Murray Darling; hh = household

Other terms used that may be considered as consistent with indicators of healthy rivers included ‘river length with healthy native vegetation’, ‘healthy flooded vegetation’, ‘population level of native fish species’, and ‘population level of native water bird and native animal species’. While vegetation, fish and water bird populations may be indicators of healthy river conditions, they may only reflect sub-sets of values for the environmental good. As a consequence, extension of the meta-analysis to encompass varying indicators of healthy rivers may be associated with scope issues. To address this in the study, dummy variables were identified for definitions that may be more narrowly scoped.

Some studies were excluded where the environmental good of interest related to river health, but where the environmental good was more broadly defined than river systems. For example Morrison et al. (2002) and Windle and Rolfe (2004) report protection values for wetlands and estuaries respectively. Other studies of river health in Australia were excluded because environmental changes were described in more qualitative terms, such as ‘poor, moderate, good and very good’ environmental health of waterways (e.g., Brouwer et al. 2010) and these could not be converted to a dependent variable consistent with other studies.

The dependent variable was described in terms of willingness to pay per household per kilometer of waterways in good health. Value estimates from many of the selected studies were not immediately comparable because of differences in attribute description, payment streams, and study year. Three key steps were required to transform values from the individual case studies into a consistent estimate of WTP per kilometer of waterways in good health. First, to address description differences, values for percentage changes were transformed into absolute values by multiplying percentage changes by river length. Second, to address variations in payment streams, all WTP estimates were converted to lump sum amounts. About 50 percent of the value estimates in the study were for lump sum values, while the remainder used annual payment streams for between three and 20 years. While there is evidence that choices are sensitive to temporal differences (Kim and Haab 2009, Taylor et al. 2003, Swait et al. 2004, Brouwer et al. 2008), little information exists to identify an appropriate discount rate. Respondents making choices where costs (and benefits) occur over long time periods are likely to have higher effective discount rates than government bonds or bank interest rates because of uncertainty about scenario outcomes and the incidence of payment burdens in the longer term. A discount rate of 15 percent has been used throughout the study to reflect this, with additional sensitivity tests at different discount rates between five percent and 25 percent not generating any significant changes in results.

Third, as the studies had been collected over a ten year period between 2000 and 2010, WTP estimates needed to be converted into real values in a consistent year. To achieve this, the Consumer Price Index (CPI) for Australia was used to bring all payment estimates into 2010 dollar equivalents. Nine values from Morrison and Bennett (2004) were identified in SPSS as extreme values, being more than three quartiles away from the mean. These estimates were omitted from the data set to give a final sample size of 145 observations. The resulting values are shown below in Figure 1.

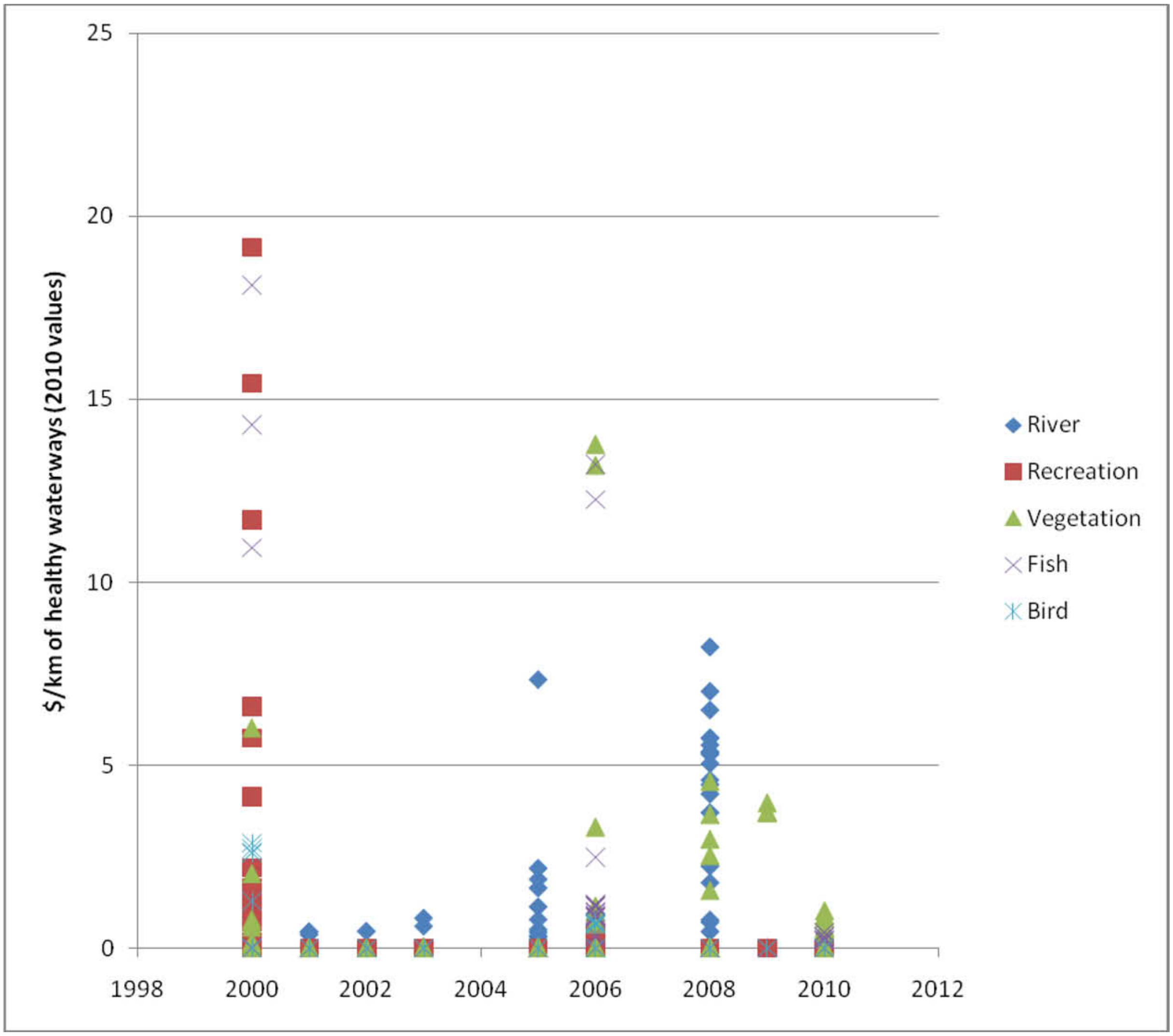


**Figure 1: Scatter plot of WTP per Km of Healthy Waterways (2010 values)**

## 4 Explanatory design variables

A number of independent variables to identify methodological, framing and design variations between the choice modelling experiments were collated across the studies, and are reviewed below, as well as being summarized in Table 2. WTP has been defined as dollars per household per kilometer of river in good health at a 15 percent discount rate.

## Table 2. Implicit prices by different good, site and study subsets

| Independent variables | Specification of variable | No. of values | Mean WTP | St. dev. | Spear. rho | MW-Z | *p* |
|---|---|---|---|---|---|---|---|
| All studies | | 145 | 2.49 | 3.76 | | | |
| ***General*** | | | | | | | |
| Year of study | Number of years from 2000 | 145 | | | 0.06 | | 0.445 |
| Refereed publication | Dummy if study published in book or refereed journal | 44 | 4.47 | 5.70 | | -2.38 | 0.017 |
| Sample size | Respondents in sample | 145 | | | 0.04 | | 0.605 |
| Mail survey | Dummy if data collected by mail survey | 66 | 3.23 | 4.97 | | -1.20 | 0.231 |
| Response rate | Response rate (percent) | 145 | | | -0.13 | | 0.109 |
| ***Amenity specification*** | | | | | | | |
| River health | Dummy if defined as general river health | 51 | 2.10 | 2.44 | | -0.51 | 0.611 |
| Recreation | Dummy if defined as health for recreation | 16 | 4.39 | 5.99 | | -1.00 | 0.316 |
| Vegetation | Dummy if defined as health for vegetation | 40 | 2.11 | 3.04 | | -0.21 | 0.834 |
| Native Fish | Dummy if defined as health for native fish | 28 | 3.01 | 5.24 | | -0.58 | 0.560 |
| Water birds | Dummy if defined as health for water birds | 10 | 1.49 | 1.10 | | -0.26 | 0.797 |
| ***Population differences*** | | | | | | | |
| Murray Darling | Dummy if in Murray Darling basin | 47 | 2.77 | 4.60 | | -0.96 | 0.337 |
| Queensland | Dummy if in Queensland | 34 | 0.78 | 1.28 | | -3.89 | 0.001 |
| Local catchment populations | Dummy if local population surveyed | 52 | 3.49 | 4.51 | | -2.82 | 0.005 |
| State capitals | Dummy if state capital population surveyed | 37 | 1.96 | 3.22 | | -1.07 | 0.284 |
| Percent male | Percent of male respondents in study | 145 | | | -0.01 | | 0.941 |
| Age | Mean age of respondents in study | 145 | | | 0.22 | | 0.008 |
| Income | Mean income of households in study | 145 | | | 0.19 | | 0.020 |
| ***Framing of tradeoffs*** | | | | | | | |
| Framed as absolute changes | Dummy if CM levels reflect actual amounts | 128 | 2.66 | 3.92 | | -1.25 | 0.212 |
| Framed as marginal changes | Dummy if CM levels are changes from current or future base | 6 | 0.76 | 0.63 | | -0.81 | 0.416 |
| River length | Length of river in catchment | 145 | | | -0.48 | | 0.001 |
| Future base lower than current | Dummy if future base lower (condition declining) | 74 | 3.30 | 4.13 | | -3.52 | 0.001 |
| Current % in good condition | % of river currently in good condition | 145 | | | -0.56 | | 0.001 |
| % of river that can be improved | Level range as % of total river length | 145 | | | -0.13 | | 0.114 |
| Year of benefit not specified | Dummy if time frame for benefits to occur not specified | 41 | 3.40 | 5.23 | | -0.44 | 0.664 |
| ***Payment mechanisms*** | | | | | | | |
| Annual payments | Dummy if mechanism is annual payments | 85 | 1.43 | 2.07 | | -5.26 | 0.001 |
| Rate or levy | Dummy if vehicle is rate or levy | 62 | 2.80 | 4.44 | | -0.51 | 0.609 |
| Mixed mechanisms | Dummy if vehicle is combination of payment types | 55 | 1.99 | 2.38 | | -0.92 | 0.357 |
| ***Presentation of levels*** | | | | | | | |
| Absolute levels only | Dummy if absolute levels | 76 | 2.60 | 3.28 | | -1.47 | 0.142 |
| Percent levels only | Dummy if percent levels | 52 | 1.22 | 1.61 | | -3.28 | 0.001 |
| Symbol levels | Dummy if symbol levels | 43 | 3.43 | 5.04 | | -2.31 | 0.021 |
| Mixed formats | Dummy if different presentation mechanisms used | 26 | 1.85 | 1.78 | | -0.56 | 0.574 |
| ***Choice set dimensions*** | | | | | | | |
| Labeled alternatives | Dummy if labels used for alternatives | 8 | 0.99 | 0.79 | | -0.73 | 0.467 |
| Number of choice cards | Number of choice sets in experiment | 145 | | | 0.03 | | 0.688 |
| Number of alternatives | Number of alternatives in each choice set | 145 | | | -0.06 | | 0.469 |
| Number of attributes | Number of attributes in each choice set | 145 | | | -0.13 | | 0.114 |
| ***Statistical models*** | | | | | | | |
| Conditional logit (CL) | Dummy if analysis uses CL/MNL | 61 | 1.65 | 2.19 | | -1.80 | 0.072 |
| Nested logit | Dummy if analysis uses Nested Logit | 59 | 3.53 | 5.18 | | -1.55 | 0.121 |
| Random parameters logit (RPL) | Dummy if analysis uses RPL | 18 | 2.68 | 1.84 | | -1.62 | 0.105 |
| Adjusted $\rho^2$ | Model fit in terms of adjusted rho-square | 145 | | | 0.10 | | 0.245 |

Specification of the amenity to be valued varied in two key ways. Variations in definition, as explained earlier, meant that different amenity scopes may have been involved across case studies. Some scopes, such as for vegetation and water birds, are likely to comprise of largely non-use values, while other scopes, such as for recreation, are likely to be focused on use values. These values are likely to be components of total use values, and hence will be subsets of values for healthy waterways as a whole. Amenity specification is also likely to vary with catchment characteristics, where factors such as size (river length), location (State) and type (inland versus coastal) may influence how respondents view the tradeoffs. Many studies could be identified in two major catchments: the Murray Darling river system draining parts of Queensland, New South Wales, Victoria and South Australia, and the Fitzroy River system in central Queensland.

Values may also vary across populations and with population characteristics. There is some evidence from individual case studies that values differ according to whether the population sample comes from inside or outside catchments (van Bueren and Bennett 2004, Morrison and Bennett 2004, Kragt et al. 2007, Bennett et al. 2008a, Bennett et al. 2008b), and when state capital versus local populations are sampled (Rolfe et al. 2002, Morrison and Bennett 2004, Kragt et al. 2007, Bennett et al. 2008a, Bennett et al. 2008b, Mazur and Bennett 2009). There is also consistent evidence across the studies that key socio-demographic characteristics such as age, gender and household income influence WTP amounts.

Several differences were identified between the studies in terms of the way that the tradeoffs were framed to choice respondents. All the experiments were consistent in terms of presenting a status quo or constant base option plus two or more improvement options together with a cost attribute. Most studies presented the information in absolute terms (kilometers of healthy waterways under different policy options), but one study (van Bueren and Bennett 2004) framed the information as marginal changes, and one study (Windle and Rolfe 2006) presented both absolute and marginal levels together.

Differences in WTP per kilometer of improvement may also be driven by marginal effects. The total length of river systems that were assessed varied from 209,118 kilometers (Australian total) to 44 kilometers (Moorabool River), while the percentages in current good condition ranged from a low of about five percent for the Clarence River (Morrison and Bennett 2004) and the Goulburn River to 65 percent for the Georges River (Kragt and Bennett 2009a, Kragt and Bennett 2009b). It is expected that respondents considered this information when indentifying their values per each one kilometer improvement.

There were differences in the way that condition trends were depicted, with the future base lower than current conditions in 57 percent of the experiments and equal to current conditions in the remainder. Concerns about losses, in the form of endowment effect, may mean that choice behavior is different between the framing scenarios. There were also differences in the total range of improvement levels offered, from a low of two percent of total river length (Mazur and Bennett 2009, 2010) to a high of 100 percent of total river (Morrison and Bennett 2004). Where the proportion amounts of level changes are higher, respondents may find improvement options more attractive.

A number of different payment mechanisms have been applied in the different studies, with most using some form of local rates or levies to identify how payment would be collected. Some studies present a mix of payment vehicles, where respondents were informed that the higher costs would be generated by a mix of higher taxes, rates, charges and consumer costs. About half of the studies involved annual costs over a number of years, with 20 year time frames being the most common.

There was some variation in survey collection techniques, with 53 percent of samples collected through mail surveys, and 47 percent collected through drop-off/pick-up techniques. The mean sample size was 378 respondents (standard deviation = 587), while the mean response rate was 41.5 percent (standard deviation =17.4 percent). Response rates were significantly lower with mail surveys, with average response rates of 33.3 percent for mail surveys and 48.8 percent for other collection methods.

Differences were also identified between studies in the way that levels were presented in the choice sets. Tradeoffs were described in absolute numbers (i.e. kilometers of waterways) in 36 percent of the experiments, in percentage terms in 10 percent of the experiments, and with the use of symbols in 38 percent of the surveys. Other formats included the joint use of absolute numbers and percentage terms (15 percent of the surveys) and the joint use of absolute numbers and symbols (six percent of the surveys).

There was limited variation in the dimensions of choice sets used in the experiments. All experiments used three alternatives (including one as a base), apart from one experiment which had five choice alternatives. The latter was also the only labeled experiment. Five attributes were used in 82 percent of the experiments, with four attributes used in the remainder. Five choice tasks per experiment were applied in 72 percent of the studies, with six choice tasks in 18 percent and eight choice tasks in nine percent.

The statistical models used in the data analysis were generally confined to three main approaches when only models used to predict results were considered. Conditional logit models were employed for 38 percent of the studies, nested logit models for 52 percent of the studies, and random parameters logit models were used for nine percent of the studies. Reported model fits in terms of adjusted rho-square values ranged from a low of 0.03 to a high of 0.41. Forty-one percent of the studies had been published in refereed journal articles or book chapters, while the remainder of the studies as conference papers and research reports.

# 5 Bivariate Analysis

The relationships between implicit prices (WTP per kilometer of healthy waterways) and the potential explanatory variables are first explored using bivariate analysis. Results are summarized in Table 2, showing average implicit prices for subsets of the data defined by different environmental attributes and other study design characteristics. Where independent variables are metric rather than grouping variables, the coefficient of correlation between the variable and implicit prices are reported. Tests for significance have been performed between different groups using the non-parametric Mann-Whitney test for independent samples and the non-parametric Spearman correlation test.

Starting with general study characteristics, there are higher values associated with refereed publications, indicating that the use of grey literature in benefit transfer studies may not generate undue biases. Correlation tests revealed a negative relationship with higher response rates, suggesting some level of self selection bias exists in samples with lower response rates. No significant effect can be detected between WTP values and years since 2000 or data collection methods. The latter showed that mail surveys tended to generate higher values, but the difference with the overall average is not statistically significant. However, the use of mail surveys was more highly correlated with annual payment mechanisms (r = -0.69) and response rates (r = -0.54), as well as other design factors.

The results for the amenity specification tests show that WTP values for recreation, vegetation and fish focused river health values are not significantly different than when only river health values were assessed, perhaps because respondents treat the indicators as proxies for the health of the river system or do not treat overall river health as a more encompassing good. These results suggest that there may be some form of amenity mis-specification involved, as values for component assets would normally be expected to be significantly lower than values for the whole asset. The relatively high values for recreation focused waterways relative to environmental asset specifications suggest that a large component of protection values relate to use values. Values are only significantly lower for experiments valuing improved river health in Queensland, an expected result given the larger areas and better conditions of the assets.

The population subgroups show that values are significantly larger for studies conducted for local or within-catchment populations, and lower for studies assessing values from capital city populations, but the latter are not significantly different from the overall mean. The significant positive correlation coefficients for average age and income show that values tend to increase with larger levels of those variables. Results of the framing comparisons show that values do not differ significantly when results are framed as marginal changes or in absolute amounts. Values are significantly higher, however, when the future base is lower than current conditions, indicating that respondents are more concerned when river health is declining. This suggests at the same time that WTP to avoid a welfare loss (equivalent variation) is higher than WTP to secure a welfare gain (compensating variation), corresponding with the findings in Barrio and Loureiro (2010).

The correlation tests show that WTP values tend to be significantly lower when larger rivers in the catchment are involved, suggesting decreasing marginal utility. No significant effect can be detected when potential improvements represented by the range of levels in the choice sets are proportionally larger.

Payment mechanisms are an important influence on implicit prices, with lump sum payments generating significantly higher implicit prices and regular payment streams generating lower implicit prices. The use of rates or levies generates higher values while the use of mixed payment mechanisms (where respondents are told that they would pay higher costs through a combination of different mechanisms) were associated with lower WTP values than the overall average, but these differences are not statistically significant.

Information communication in choice sets through the way that levels are presented appears to have a large influence on results. Values are higher when levels are presented as symbols, and lower when levels are presented as percentage values. Correlation tests with choice set dimensions show that increasing the number of choice tasks is generally associated with lower values, while increasing the number of choice alternatives and attributes is associated with higher values, but these differences are not significant.

The groupings associated with statistical analysis methods indicate that conditional logit models are associated with significantly lower values (average value is 33 percent lower). Values from mixed logit (RPL) models were slightly (eight percent) higher than the average, but the difference was only significant at the 11 percent level.

# 6 Meta-analysis regression model

The estimates for the regression coefficients $\beta_j$ in equations (2) to (4) are obtained through maximum likelihood (ML) techniques. The dependent variable used is the log of the WTP per kilometre of river health, and the independent variables used are described in Table 2. The tobit model essentially identifies how the different independent variables (including both dummy and metric variables) influence WTP. The results of the modelling are shown in Table 3.

**Table 3: Tobit meta-regression model explaining WTP for river health**

| Dependent variable = LN of $/km | Coefficient estimate | Standard error |
|---|---|---|
| Constant | 12.071*** | 4.643 |
| Year of study (from 2000) | -0.284** | 0.144 |
| ***Population differences*** | | |
| Income ($000) | 0.031*** | 0.007 |
| ***Framing of tradeoffs*** | | |
| Future base lower than current | 0.939*** | 0.243 |
| Level range as % of river length | -0.022** | 0.011 |
| Years of benefit not specified | 1.650*** | 0.597 |
| ***Payment mechanism*** | | |
| Annual payments | -2.830*** | 1.135 |
| Rate/levy payment vehicle | -2.667*** | 1.006 |
| ***Data collection method*** | | |
| Mail survey | -2.078* | 1.201 |
| Response rate | 0.035*** | 0.013 |
| ***Presentation of levels*** | | |
| Levels in percentages | -1.554*** | 0.461 |
| Levels as symbols | 1.737*** | 0.622 |
| ***Choice set dimensions*** | | |
| Number of choice tasks | -0.686* | 0.360 |
| Number of alternatives | -0.234 | 0.314 |
| Number of attributes | -1.340** | 0.699 |
| ***Model summary statistics*** | | |
| Standard deviation random effects ($\sigma_u$) | $0.994 \times 10^{-11}$ | 0.149 |
| Standard deviation fixed effects ($\sigma_\varepsilon$) | 0.895*** | 0.085 |
| Log-likelihood | -112.152 | |
| Chi square statistic (14 degrees of freedom) | 100.29*** | |
| Number of observations | 145 | |

$^{*}p<0.10$; $^{**}p<0.05$; $^{***}p<0.01$.

The small data set limited the number of attributes that could be identified as significant. Amenity and population differences were identified in relation to the year of the study (values declining from 2000), and the average household income of samples (positively related to values). In the regression model, 145 observations were retrieved from the 19 studies presented in the previous section. This implies, on average, seven to eight observations from each individual study. We accounted for possible clustering of these multiple values from single studies, i.e., heteroskedasticity caused by the use of a similar design in the studies from which these values originate, in the random component of our model. However, we were unable to find a significant random effect. The Likelihood Ratio test is unable to reject the null hypothesis that the standard error of the random effects parameter associated with study ($\sigma_u$) equals zero, possibly due to the inclusion of a wide variety of control variables for the different design characteristics in the fixed part of the model.

As expected, the use of compulsory rates or levies for payment mechanisms was significant in lowering values. However WTP through annual payment mechanisms were identified as significantly lower than WTP through lump sum payments. This indicates that the discount rate used in this analysis (15 percent) to convert period payments to lump sum amounts is too high to drive equivalence. However, these differences in WTP remain even when discount rates as low as five percent are used, suggesting that further work is necessary to understand how respondents view payment vehicles and internally discount future payments. Case studies where the years of environmental improvement were not specified are associated with higher WTP.

Only some factors relating to the design and analysis of the choice experiments could be identified as significant. Increasing the number of choice tasks and the number of attributes in an experiment appears to reduce WTP estimates. The latter effect could be caused by omitted variable bias where smaller numbers of less precisely defined attributes lead to over estimates of values. Both effects could also be caused by some level of complexity or fatigue effects, although these would be more likely to emerge in the random terms rather than coefficient estimates. In contrast, the number of alternatives was not significant, most probably due to the limited variation between three and five alternatives across the data set. The way that tradeoffs were framed was also identified as having a significant influence on values. Having declining conditions (as shown by a lower future base) increased WTP, while increasing marginal improvements (level range relative to river length) were associated with lower marginal WTP. The way survey respondents are informed about hypothetical changes and levels are communicated also plays a significant role. The description of levels in percentage terms reduced values, while presentation in symbols increased them (compared to the baseline of mainly absolute levels).

Finally, mail surveys were identified as a data collection mechanism that led to lower values, while values were positively correlated with response rates.

## 7 Conclusions

The literature on choice modeling to value environmental change is rapidly increasing. Although design factors are tested to different degrees in individual studies, variations in case studies, framing and methodological applications make it difficult to compare design factors across studies and to identify their relative influence. In this review, we aimed to satisfy the need for research synthesis through the use of a statistical meta-analysis to aggregate information and insights based on experiences in the specific domain of river health in Australia. Such meta-analysis of the role of design factors in choice experiments is currently absent in the literature.

The results of this study show substantial variation in the WTP of households for river health in Australia, although there is some indication that values are declining and exhibit less variation over time (Figure 1). There is also some evidence of amenity misspecification, where values for overall river health are lower than for components such as fish and recreation, suggesting that use values dominate over nonuse values. Differences are, however, not statistically significant, also not when controlling for other influencing factors in the multivariate meta-regression model.

The results of both the bi-variate and multi-variate analysis demonstrate that several design factors appear to have a systematic effect on estimates of WTP. Values are particularly sensitive to the way that tradeoffs are framed, where the relative size of marginal changes are important. There is also evidence that values are sensitive to a number of design issues around the collection of data, representation of levels to respondents in choice experiments and choice set dimensions such as number of choice tasks and attribute levels. These results are consistent with a number of studies such as Louviere et al. (2000), Breffle and Rowe (2002), Caussade et al. (2005), Hensher (2006), Rolfe and Bennett (2009). However, the impact of the type of model used in the analysis of results found in the bivariate analysis could not be replicated in the multivariate regression analysis, indicating that much of the attention paid to methods of statistical analysis may have limited impact on value estimates and policy applications.

There is some evidence that mail surveys generate lower WTP, and that some form of complexity or fatigue effects may be associated with experiments that have more choice tasks and more attribute levels, consistent with Louviere et al. (2000), Caussade et al. (2005) and Hensher (2006). Values also appear to be sensitive to the way that the payment mechanism and frequency is structured, a design dimension receiving little attention in stated preference experiments.

These results indicate that choice experiment design in Australia may have important significant impacts on values. However, issues of amenity misspecification, tradeoff framing and payment mechanisms that were identified in this study as having systematic influences on values are rarely the subject of methodological testing in the existing literature. The results of this meta-analysis study suggest that more attention should be paid to the way that tradeoffs are framed in choice experiments relative to internal choice set structure and data analysis.

# 8 References

Barrio, M. and M. L., Loureiro, 2010. A meta-analysis of contingent valuation forest studies. *Ecological Economics,* 69(5), 1023-1030.

Bateman, I. J. and A. P., Jones, 2003. Contrasting conventional with multi-level modelling approaches to meta-analysis: expectation consistency in UK woodland recreation values. *Land Economics*, 79(2), 235-258.

Bateman, I. J., R., Brouwer, S., Ferrini, M., Schaafsma, D.N., Barton, A., Dubgaard, B., Hasler, S., Hime, I., Liekens and S., Navrud, 2011. Making benefit transfers work: deriving and testing principles for value transfers for similar and dissimilar sites using a case study of the non-market benefits of water quality improvements across Europe. *Environmental and Resource Economics*, 50(3), 365-387.

Bennett, J. W., R., Dumsday, G., Howell, C., Lloyd, N., Sturgess and L., Van Raalte, 2008a. The economic value of improved environmental health in Victorian Rivers. *Australasian Journal of Environmental Management*, 15(3), 138-148.

Bennett, J. W., R., Dumsday and R., Gillespie, 2008b. Analysing options for the River Red Gum forests along the Murray River. Paper presented at 52nd Australian Agricultural and Resource Economics Society conference, Queenstown NZ.

Bergstrom, J. C. and L. O., Taylor, 2006. Using meta-analysis for benefits transfer: theory and practice. *Ecological Economics*, 60(2), 351-360.

Brander, L. M., R. J. G. M.l Florax and J. Vermaat, 2006. The empirics of wetland valuation: A comprehensive summary and a meta-Analysis of the literature. *Environmental and Resource Economics,* 33(2), 223-250.

Breffle, W. S. and R. D., Rowe, 2002. Choice Question Formats and Natural Resource Tradeoffs. *Land Economics*, 78(2), 298-314.

Brookshire, D. S. and H. R., Neill, 1992. Benefit transfers: conceptual and ethical issues. *Water Resources Research*, 28(3), 651-655.

Brouwer, R. and F. A., Spaninks, 1999. The validity of environmental benefits transfer: further empirical testing. *Environmental and Resource Economics,* 14(1), 95-117.

Brouwer, R., 2000. Environmental value transfer: State of the art and future prospects. *Ecological Economics,* 32(1), 137-152.

Brouwer, R., I. H., Langford, I. J., Bateman and R. K., Turner, 1999. A meta-analysis of wetland contingent valuation studies. *Regional Environmental Change*, 1(1), 47-57.

Brouwer, R. and I. J., Bateman, 2005. Benefits transfer of willingness to pay estimates and functions for health-risk reductions: A cross-country study. *Journal of Health Economics,* 24(3), 591-611.

Brouwer, R., P., van Beukering and E., Sultanian, 2008. The impact of the bird flu on public willingness to pay for the protection of migratory birds. *Ecological Economics,* 64(3), 575-585.

Brouwer, R., 2009. *Multi-attribute choice modeling of Australia's rivers and wetlands: A meta-analysis of 10 years of research*. CSIRO Working Paper 2009-05, ISSN 1834-5638.

Brouwer, R., T. Dekker, J. C. Rolfe and J. Windle, 2010. Choice Certainty and Consistency in Repeated Choice Experiments. *Environmental and Resource Economics,* 46(1), 93-109.

Caussade, S., J.deD. Ortuzar, L., Rizzi and D., Hensher, 2005. Assessing the Influence of Design Dimensions on Stated Choice Experiment Estimates. *Transportation Research Part B*, 39(7), 621-640.

Glass, G. V., B., McGaw and M. L., Smith, 1981. *Meta-analysis in social research*. Sage Publications, Beverly Hills.

Hatton MacDonald, D., M., Morrison, J. M., Rose and K., Boyle, 2010. Untangling differences in values from internet and mail stated preference studies. paper presented at the 4th World Congress on Environmental Economics, Montreal, Canada, June 28 – July 2.

Hensher, D. A., 2006. Revealing Differences in Willingness to Pay due to the Dimensionalities of Stated Choice Designs, an Initial Assessment. *Environmental and Resource Economics,* 34(1), 7-44.

Johnston, R. J., E. Y., Besedin, R., Iovanna, C., Miller, R., Wardwell and M., Ranson, 2005. Systematic variation in willingness to pay for aquatic resource improvements and implications for benefit transfer: A Meta-Analysis. *Canadian Journal of Agricultural Economics*, 53(2-3), 221-248.

Johnston, R. J, E. Y., Besedin and R. F., Wardwell, 2003. Modeling relationships between use and nonuse values for surface water quality: a meta-analysis. *Water Resources Research,* 39(12), 1363-1372.

Johnston, R .J. and R. S., Rosenberger, 2010. Methods, trends and controversies in contemporary benefit transfer. *Journal of Economic Surveys*, 24(3), 479-510.

Kim, S.-I. and T. C., Haab (2009) Temporal sensitivity of willingness to pay and implied discount rates. *Resource and Energy Economic*, 31(2), 89-102.

Kragt, M., J.W., Bennett, C., Lloyd and R., Dumsday, 2007. Comparing choice models of river health improvement for the Goulburn River. paper presented at 51st AARES conference, 13-16 February 2007 Queenstown, New Zealand.

Kragt, M. and J. W., Bennett, 2009a. Using choice experiments to value river and estuary health in Tasmania with individual preference heterogeneity. Paper presented at 53rd Australian Agricultural and Resource Economics Society conference, Cairns.

Kragt, M. and J. W., Bennett, 2009b. *Using choice experiments to value river and estuary health in Tasmania with individual preference heterogeneity*. Environmental Economics Research Hub Research Report 16, Australian National University, Canberra.

Kragt, M. and J. W., Bennett, 2010. The impacts of attribute level framing and changing cost levels on choice experiments value estimates. Paper presented at 54th Australian Agricultural and Resource Economics Society conference, Adelaide.

Lindhjem, H., 2007. 20 years of stated preference valuation of non-timber benefits from Fennoscandian forests: A meta-analysis. *Journal of Forest Economics,* 12(4), 251-277.

Louviere, J., D., Hensher and J., Swait, 2000. *Stated Choice Models - Analysis and Application*. Cambridge University Press, Cambridge, U.K.

Mazur, K. and J. W., Bennett, 2009. Location differences in communities' preferences for environmental improvements in selected NSW catchments: A choice modelling approach. Paper presented at 53rd Australian Agricultural and Resource Economics Society conference, Cairns.

Mazur, K. and J. W., Bennett, 2010. Framing for incentive compatibility in choice modelling. Paper presented at 54th Australian Agricultural and Resource Economics Society conference, Adelaide.

Mazur, K. M. (2011) *Choice Modelling in the Development of Natural Resource Management Strategies in NSW Catchments*. PhD thesis, Australian National University, Canberra.

Morrison, M., J. W., Bennett, R. K., Blamey and J. J., Louviere, 2002. Choice modelling and tests of benefit transfer. *American Journal of Agricultural Economics*, 84(1), 161-170.

Morrison, M. and J. W. Bennett, 2004. Valuing New South Wales rivers for use in benefits transfer. *Australian Journal of Agricultural and Resource Economics,* 48 (4), 591-611.

Morrison, M. and J., Bennett, 2006. Valuing New South Wales rivers for use in benefits transfer, In: Rolfe, J. and Bennett, J. (eds) *Choice modelling and the transfer of environmental values*. Edward Elgar Publishing, Cheltenham, UK.

Morrison, M., D., Hatton MacDonald, K., Boyle and J. M., Rose, 2010. Ecological values for a major river system at risk, unpublished paper reported in M. Morrison and D. Hatton MacDonald, Economic Valuation of Environmental Benefits in the Murray-Darling Basin, report prepared for the Murray Darling Basin Authority, Canberra.

Nelson, J. P. and P. E., Kennedy, 2009. The use (and abuse) of meta-analysis in environmental and resource economics: an assessment. *Environmental and Resource Economics*, 42(3), 345-377.

Rolfe, J. (2006) Theoretical issues in using choice modelling data for benefit transfer. in J. C. Rolfe and J. W., Bennett (eds) 2006. *Choice Modelling and the Transfer of Environmental Values*. Edward Elgar, Cheltenham, UK.

Rolfe, J. C. and J. W., Bennett (eds) 2006. *Choice Modelling and the Transfer of Environmental Values*. Edward Elgar, Cheltenham, UK.

Rolfe, J. C. and J. W., Bennett, 2009. The impact of offering two versus three alternatives in Choice Modelling experiments. *Ecological Economics*, 68(4), 1140-1148.
Rolfe, J. C., A., Loch and J. W., Bennett, 2002. *Tests of benefits transfer across sites and population in the Fitzroy basin*. Valuing floodplain development in the Fitzroy basin Research Report no. 4, Central Queensland University, Rockhampton.
Rolfe, J. C. and J., Windle, 2003. Valuing the protection of aboriginal cultural heritage sites. *The Economic Record*, 79(Special issue), s85 – s95.
Rosenberger, R. S. and J., Loomis, 2000. Using meta-analysis for benefit transfer: in-sample convergent validity tests for an outdoor recreation database. *Water Resources Research,* 36(4), 1097-1107.
Rosenberger, R. S. and T. D., Stanley, 2006. Measurement, generalization and publication: sources of error in benefit transfers and their management. *Ecological Economics*, 60(2), 372-378.
Smith, V. K. and S. K., Pattanayak, 2002. Is meta-analysis a Noah's Ark for non-market valuation?. *Environmental and Resource Economics*, 22(1-2), 271-296.
Swait, J., V. L., Adamowicz and M., van Bueren, 2004. Choice and Temporal Welfare Impacts: Incorporating History into Discrete Choice Models. *Journal of Environmental Economics and Management*, 47(1), 94-116.
Taylor, C. L., V. L., Adamowicz and M. K., Luckert, 2003. Preferences over the Timing of Forest Resource Use. *Journal of Forest Economics*, 9(3), 223-240
Van Bueren, M. and J. W., Bennett, 2004. Towards the development of a transferable set of value estimates for environmental attributes. *Australian Journal of Agricultural and Resource Economics,* 48(1), 1-32.
Van den Bergh, J. C .J., M., K. J., Button, P., Nijkamp and G. C., Pepping, 1997. *Meta-analysis in Environmental Economics.* Kluwer, Dordrecht, The Netherlands.
Van Houtven, G., J., Powers and S. K., Pattanayak, 2007. Valuing water quality improvements in the United States using meta-analysis: is the glass half-full or half-empty for national policy analysis? *Resource and Energy Economics*, 29(3), 206-228.
Windle, J. and J. C., Rolfe, 2004. *Assessing values for estuary protection with choice modelling using different payment mechanisms*. Valuing floodplain development in the Fitzroy basin Research Report no.10.
Windle, J. and J. C., Rolfe, 2006. *Non market values for improved NRM outcomes in Queensland.* Research report 2 in the non-market valuation component of AGSIP project # 13, CQUniversity, Rockhampton.
Woodward, R. T. and Y.-S., Wui, 2001. The economic value of wetland services: A meta-analysis. *Ecological Economics,* 37(2), 257–270.
Zandersen, M. and R. S. J., Tol, 2009. A meta-analysis of forest recreation values in Europe. *Journal of Forest Economics,* 15(1-2), 109-130.